BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 5: e74178
doi: 10.3897/biss.5.74178

OPEN ACCESS

Conference Abstract

The Fungal Literature-based Occurrence Database in Southern West Siberia (Russia)

Nina Filippova[‡], Dmitry Ageev[§], Sergey Bolshakov[|], Olga Vayshlya[¶], Anastasia Vlasenko[#], Vyacheslav Vlasenko[#], Sergei Gashkov[¶], Irina Gorbunova[#], Eugene A. Davydov[#], Elena Zvyagina[¤], Nadezhda Kudashova[¶], Maria Tomoshevich[#], Aleksandra Filippova[«], Natalia Shabanova[¶], Lidia Yakovchenko[#], Irina Vorob'eva[#], Ludmila Kalinina[|], Ekaterina Palomozhnykh[|]

‡ Yugra State University, Khanty-Mansiysk, Russia
§ OOO (Limited Liability Company) "SIGNATEC", Novosibirsk, Russia
| Komarov Botanical Institute of the Russian Academy of Sciences, Saint Petersburg, Russia
¶ National Research Tomsk State University, Tomsk, Russia
# Central Siberian Botanical Garden, Novosibirsk, Russia
¤ Lomonosov Moscow State University, Moscow, Russia
« Kemerovo State University, Kemerovo, Russia

Corresponding author: Nina Filippova (filippova.courlee.nina@gmail.com)

Received: 08 Sep 2021 | Published: 10 Sep 2021

Citation: Filippova N, Ageev D, Bolshakov S, Vayshlya O, Vlasenko A, Vlasenko V, Gashkov S, Gorbunova I, Davydov EA, Zvyagina E, Kudashova N, Tomoshevich M, Filippova A, Shabanova N, Yakovchenko L, Vorob'eva I, Kalinina L, Palomozhnykh E (2021) The Fungal Literature-based Occurrence Database in Southern West Siberia (Russia). Biodiversity Information Science and Standards 5: e74178.
https://doi.org/10.3897/biss.5.74178

Abstract

The abstract presents the initiative to develop the Fungal Literature-based Occurrence Database for Southern West Siberia (FuSWS), which mobilizes occurrences of fungi from published literature (literature-based occurrences, Darwin Core MaterialCitation). The FuSWS database includes 28 fields describing species name, publication source, herbarium number (if exists), date of sampling or observation, locality information, vegetation, substrate, and others.

The initiative on digitization of literature-based occurrence data started in the northern part of Western Siberia two years ago (Filippova et al. 2021a). The present project extends the initiative to the south and includes eight administrative regions (Sverdlovsk, Omsk, Kurgan, Tomsk, Novosibirsk, Kemerovo, Altay, and Gorny Altay). The area occupies the central to

© Filippova N et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

southern part of the West Siberian Plain. It extends for about 1.5 thousand km from the west to the east from the eastern slopes of the Ural Mountains to Yenisey River, and from north to south—about 1.3 thousand km. The total area equals about 1.2 million km^2.

Currently, the project is actively growing in spatial, collaboration and data accumulation terms. The working group of about 30 mycologists from 16 organizations dedicated to the digitization initiative was created as part of the Siberian Mycological Society (informal organization since 2019). They have created the most complete bibliographic list of mycology-related papers for the Southern West Siberia, including over 800 publications for the last two centuries (the earliest dated 1800). At abstract submission, the database had been populated with a total of about 10K records from about 100 sources. The dataset is uploaded to GBIF, where it is available for online search of species occurrences and/or download (Filippova et al. 2021b) Fig. 1. The project's page with the introduction, templates, bibliography list, video-presentations and written instructions is available at the website of the Siberian Mycological Society (https://sibmyco.org/literaturedatabase).

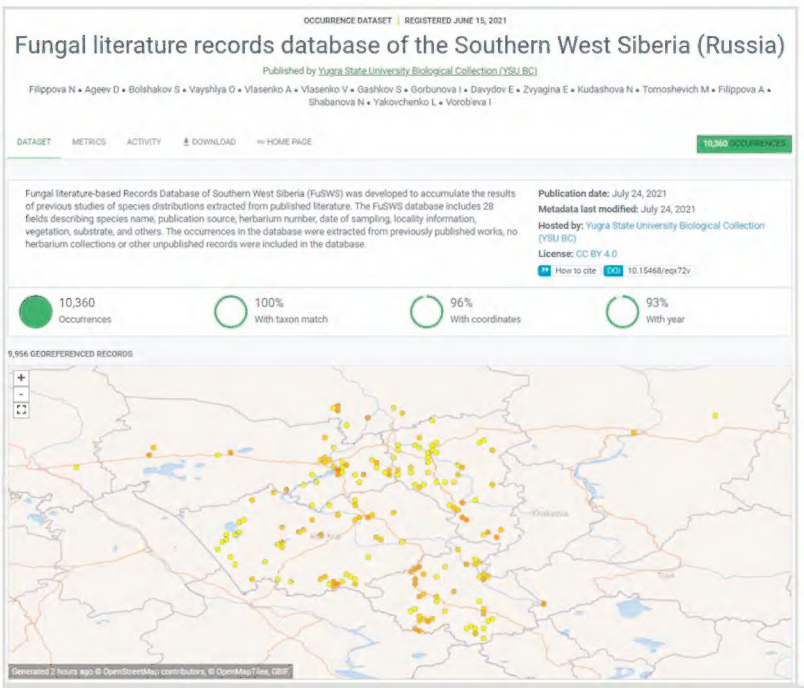
OCCURRENCE DATASET | REGISTERED JUNE 15, 2021
Fungal literature records database of the Southern West Siberia (Russia)
Published by Yugra State University Biological Collection (YSU BC)
Filippova N • Ageev D • Bolshakov S • Vayshlya O • Vlasenko A • Vlasenko V • Gashkov S • Gorbunova I • Davydov E • Zvyagina E • Kudashova N • Tomoshevich M • Filippova A • Shabanova N • Yakovchenko L • Vorobieva I
DATASET METRICS ACTIVITY DOWNLOAD HOME PAGE
10,360 OCCURRENCES
Fungal literature-based Records Database of Southern West Siberia (FuSWS) was developed to accumulate the results of previous studies of species distributions extracted from published literature. The FuSWS database includes 28 fields describing species name, publication source, herbarium number, date of sampling, locality information, vegetation, substrate, and others. The occurrences in the database were extracted from previously published works, no herbarium collections or other unpublished records were included in the database.
Publication date: July 24, 2021
Metadata last modified: July 24, 2021
Hosted by: Yugra State University Biological Collection (YSU BC)
License: CC BY 4.0
How to cite DOI 10.15468/eqx72v
10,360 Occurrences
100% With taxon match
96% With coordinates
93% With year
9,956 GEOREFERENCED RECORDS

Figure 1.

The screenshot of the dataset page with about 10K digitized literature-based records of fungi for the Southern West Siberia regions published in GBIF.

The following protocol describes the digitization workflow in detail:

1. The bibliography of related publications is compiled using Zotero bibliographic manager. Only published works (peer-reviewed papers, conference proceedings, PhD theses, monographs or book chapters) are selected. If possible, the sources are digitized and added to the library as PDF files.

2. The template of the FuSWS database is made with Google Sheets, which allows simultaneous use by several specialists, in a common data format provided. The simple Microsoft Excel template is also available for the offline databasing. The Darwin Core standard is applied to the database field structure to accommodate the relevant information extracted from the publications.

3. From the available bibliography of publications related to the region, only works with species occurrences are selected for the databasing purpose. The main source of occurrences is annotated species lists with exact localities of the records. However, different sorts of other species citations are also extracted, provided that they had the connection to any geography.

4. All occurrences are georeferenced, either from the coordinates provided in the paper, or from the verbatim description of the field work locality. The georeferencing of the verbatim descriptions is made using Yandex or Google map services. Depending on the quality of georeference provided in publications, the uncertainty is estimated as follows: 1) the coordinate of a fruiting structure or a plot provided in the publication gives the uncertainty about 3-30 meters; 2) the coordinate of the field work locality provided in publication gives the uncertainty about 500 m to 5 km; 3) the report of the species presence in a particular region gives the centroid of the area with the uncertainty radius to include its borders.

5. The locality names reported in Russian are translated to English and written in the «locality» field. Russian descriptions are reserved in the field «verbatimLocality» for accuracy.

6. When possible, the «eventDate» is extracted from the annotation data. Whenever this information is absent, the date of the publication is used instead with the remarks in the «verbatimEventDate» field.

7. The ecological features, habitat and substrate preferences are written in the «habitat» field and reserved in Russian.

8. The original scientific names reported in publications are filled in the «originalNameUsage» field. Correction of spelling errors is made using the GBIF Species Matching tool. This tool is also used to create the additional fields of taxonomic hierarchy from species to kingdom, to fill in the «taxonRank» field and to synonymize according to the GBIF Backbone Taxonomy.

9. To track the digitization process, a worksheet is maintained. Each bibliographic record has a series of fields to describe the digitization process and its results: the total number of extracted occurrence records, general description of the occurrence quality, presence of the observation date, details of georeferencing and the name of a person responsible for the digitization.

Keywords

materialCitation, fungi, digitization, biodiversity data mobilization, GBIF

Presenting author

Nina Filippova

Presented at

TDWG 2021

References

- Filippova N, Arefyev S, Bulyonkova T, Zvyagina E, Kapitonov V, Makarova T, Mukhin V, Stavishenko I, Tavshanzhi E, Shiryaev A, Tolpysheva T, Sedelnikova N, Ryabitseva N, Paukov A, Zhurbenko M (2021a) Fungal literature records database of the Northern West Siberia (Russia). Version 1.13. Yugra State University Biological Collection (YSU BC). Occurrence dataset https://doi.org/10.15468/hfje3l
- Filippova N, Ageev D, Bolshakov S, Vayshlya O, Vlasenko A, Vlasenko V, Gashkov S, Gorbunova I, Davydov E, Zvyagina E, Kudashova N, Tomoshevich M, Filippova A, Shabanova N, Yakovchenko L, Vorob'eva I (2021b) Fungal literature records database of the Southern West Siberia (Russia). Version 1.6. Yugra State University Biological Collection (YSU BC). Occurrence dataset https://doi.org/10.15468/eqx72v